ELSEVIER Stochastic Processes and their Applications 56 (1995) 57–75

stochastic processes and their applications

Limit theorems for multitype epidemics

Håkan Andersson[a], Boualem Djehiche[b,*]

[a]Department of Mathematics, Stockholm University, Box 6701, S-113 85 Stockholm, Sweden
[b]Department of Mathematics, Royal Institute of Technology, S-100 44 Stockholm, Sweden

Received January 1993; revised August 1994

Abstract

By means of the law of large numbers and the central limit theorem, we compare the spatial epidemic model proposed by Kendall with a sequence of multitype epidemics, viewed as Hilbert space-valued stochastic processes. In the limit, the fluctuation process turns out to be an infinite-dimensional Ornstein–Uhlenbeck process.

Keywords: Epidemic process; Law of large numbers; Central limit theorem; Ornstein-Uhlenbeck process

1. Introduction

Deterministic models are frequently used in order to describe various physical, chemical and biological processes involving a large parameter (the number of molecules, individuals, etc.), even if the underlying structure is fully stochastic so that a stochastic description would in fact be more accurate. The reason for making this choice is, of course, that the deterministic approximation is in general much easier to analyze than the stochastic model, and the error made is hopefully negligible for all practical purposes. The following question then arises: How can we put such a reasoning on solid grounds?

For *global* – i.e. spatially homogeneous – phenomena, this dilemma has been dealt with by Kurtz (1971) (see also Ethier and Kurtz, 1986), who showed how to obtain various solutions of ordinary differential equations as limits of pure jump Markov processes, and who also studied second-order approximations. Arnold (1981) discusses the corresponding problem for *local* models, such as processes of chemical reactions with diffusion. On a macroscopic level such models are described by partial differential equations, and the natural stochastic counterparts are the so-called space–time jump Markov processes. Kotelenez (1986, 1988) and Blount (1991, 1993), among others, have developed Arnold's ideas to obtain laws of large numbers and

*Corresponding author.

SSDI 0304-4149(94)00059-X

central limit theorems for these Markov processes. (For further references, see Oelschläger, 1989.)

Our objective is to study a local epidemic model – a spatial version of the classical Kermack–McKendrick model – proposed by Kendall (1965). The evolution of $\bar{\xi} = (\bar{x}, \bar{y})$, where $\bar{x}$ and $\bar{y}$ represent the density of susceptible and infective individuals, respectively, is governed by a system of nonlinear partial differential equations. We obtain a stochastic analogue of Kendall's model by dividing the space into N small areas, each area containing approximately ℓ individuals, and then modelling the density $\Xi^n = (X^n, Y^n)$, $n = (N, \ell)$, of susceptibles and infectives as a multitype general epidemic. We prove the following law of large numbers (provided that all infection rates converge properly):

$$\sup_{t \in [0,T]} \|\Xi^n(t) - \bar{\xi}(t)\| \to 0 \quad \text{in probability}$$

as $N, \ell \to \infty$, where $\|\cdot\|$ denotes the $\mathscr{L}^2$-norm, and the central limit theorem:

$$\sqrt{N\ell}(\Xi^n - \bar{\xi}) \to V \quad \text{weakly},$$

if $\ell = \ell(N)$ satisfies $\ell/N \to \infty$ and $\ell/N^3 \to 0$ as $N \to \infty$, where V is an infinite dimensional (distribution space valued) Ornstein–Uhlenbeck process. It is necessary to prove the law of large numbers in a function space rather than in a space of distributions, simply because the dynamics in our model is *nonlinear* and multiplication of distributions is not a well-defined operation. In order to obtain second-order approximations, we use methods based on Stroock and Varadhan's martingale characterization of Hilbert-valued diffusion processes, as developed by Holley and Stroock (1978) and by Métivier (1984).

2. Preliminaries and notation

We first recall some basic facts about the classical Sobolev spaces (cf. Adams, 1975). Let $\|\cdot\|_0$ be the usual $\mathscr{L}^2$-norm on $[0,1]$ and set

$$\|x\|_k^2 = \|x\|_0^2 + \sum_{l=1}^{k} \|D^l x\|_0^2 \quad \text{for } k \geq 1, \tag{2.1}$$

where $D^l x$ denotes the lth derivative (in the sense of distributions) of x. Next, let $\mathscr{E}$ be the space of real-valued infinitely differentiable functions on $[0,1]$, completed in the topology induced by the norms $\|\cdot\|_k$, $k \in \mathbf{N}$, and define $\mathscr{H}_k$ to be the completion of $\mathscr{E}$ relative to $\|\cdot\|_k$; in particular, $\mathscr{H}_0 = \mathscr{L}^2([0,1])$. Maurin's theorem tells us that the imbedding $\mathscr{H}_k \to \mathscr{H}_{k-1}$ is Hilbert–Schmidt for each $k \geq 1$. This means that $\sum_m \|\varphi_m\|_{k-1}^2 < \infty$ if (φ_m) is an orthonormal set in $\mathscr{H}_k$, a fact that will be used several times in the sequel.

For fixed $k \in \mathbf{N}$, let $\mathscr{H}_{-k}$ be the dual space of $\mathscr{H}_k$ with norm given by

$$\|x\|_{-k}^2 = \sum_{m=0}^{\infty} \frac{\langle x, e_m \rangle^2}{(1 + \pi^2 m^2)^k}, \tag{2.2}$$

where $(e_m)_{m=1}^{\infty}$ is the standard complete orthonormal system (CONS) in $\mathscr{H}_0$ and $\langle \cdot, \cdot \rangle$ denotes the natural pairing between a Hilbert space and its dual space. It is well known that the spaces $\mathscr{H}_{-k}$ consist of derivatives: $x \in \mathscr{H}_{-k}$ if and only if $x = \sum_{l=1}^{k} D^l x_l$ for some $\mathscr{L}^2$-functions x_l. By polarization, we may associate with each norm $\| \cdot \|_k$ an inner product $\langle \cdot ; \cdot \rangle_k$, which turns $\mathscr{H}_k$ into a (separable) Hilbert space.

Let $\mathscr{L}_1(\mathscr{H}_k, \mathscr{H}_{-k})$ be the space of nuclear (positive) operators from $\mathscr{H}_k$ to its dual. A sufficient condition for a linear operator a to be in $\mathscr{L}_1(\mathscr{H}_k, \mathscr{H}_{-k})$, is that (see Gelfand and Vilenkin, 1964) there exists a positive constant K, such that

$$\langle \varphi_m, a\varphi_m \rangle \leq K \|\varphi_m\|_{k-1}^2,$$

where $(\varphi_m)_{m=1}^{\infty}$ is a CONS in $\mathscr{H}_k$.

Finally, let $\mathscr{E}'$ be the dual of $\mathscr{E}$, equipped with the strong topology. Then, if we identify $\mathscr{H}_0$ with its dual $\mathscr{H}_{-0}$, we obtain the scheme

$$\mathscr{E}' \supseteq \cdots \supseteq \mathscr{H}_{-1} \supseteq \mathscr{H}_0 \supseteq \mathscr{H}_1 \supseteq \cdots \supseteq \mathscr{E}.$$

The following family of finite-dimensional subspaces of $\mathscr{H}_0$ will be of fundamental importance in the sequel. For fixed $N \in \mathbf{N}$, $N \geq 1$, set

$$I_1 = [0, 1/N],$$

$$I_i = ((i-1)/N, i/N], \quad 2 \leq i \leq N,$$

and define the orthogonal (although not orthonormal!) set of indicator functions $h^i = \mathbf{1}_{I_i}$, $1 \leq i \leq N$. The functions h^i span an N-dimensional space $\mathscr{S}_N$ of step functions, whose members may conveniently be written as $x = \sum_{i=1}^{N} x_i h^i$. $\mathscr{S}_N$ inherits an inner product from $\mathscr{H}_0$:

$$\langle x; \tilde{x} \rangle_0 = \int_0^1 x(r)\tilde{x}(r)\,\mathrm{d}r = \frac{1}{N}\sum_{i=1}^{N} x_i \tilde{x}_i,$$

if $x, \tilde{x} \in \mathscr{S}_N$.

Let us also define a projection of $\mathscr{H}_k$ ($k \in \mathbf{N}$ fixed) onto $\mathscr{S}_N$ by taking conditional expectations: for $x \in \mathscr{H}_k$,

$$P_N x(r) = \sum_{i=1}^{N} \left\{ N \int_{I_i} x(r')\,\mathrm{d}r' \right\} h^i(r). \tag{2.3}$$

One easily checks that $\langle x; \tilde{x} \rangle_0 = \langle x; P_N \tilde{x} \rangle_0$ if $x \in \mathscr{S}_N$ and $\tilde{x} \in \mathscr{H}_0$. Note that $\|x - P_N x\| \to 0$ as $N \to \infty$, by the martingale convergence theorem. We have also the following lemma.

Lemma 2.1. *Fix $k \geq 2$. Then for every $x \in \mathscr{H}_k$,*

$$\|x - P_N x\|_{-k} \leq \frac{C}{N^2}, \tag{2.4}$$

where the constant C does not depend on N.

Proof. First note that, using the imbedding $\mathscr{H}_k \to C_b^1([0,1])$ (cf. Adams, 1975, p. 97), the mean value theorem yields, for every $\psi \in \mathscr{H}_k$,

$$\|(I - P_N)\psi\|_\infty \le \|D^1\psi\|_\infty / N. \tag{2.5}$$

Since the operator $I - P_N$ is idempotent,

$$\begin{aligned}\|x - P_N x\|_{-k}^2 &= \sum_{m=0}^{\infty} \frac{\langle (I - P_N)x, e_m\rangle^2}{(1 + \pi^2 m^2)^k} \\ &= \sum_{m=0}^{\infty} \frac{\langle (I - P_N)x, (I - P_N)e_m\rangle^2}{(1 + \pi^2 m^2)^k} \\ &\le \frac{C^2}{N^4} \sum_{m=1}^{\infty} m^{2-2k},\end{aligned}$$

by (2.5). □

In the sequel, we will actually be working in product spaces $\mathscr{H}^2$, with $\mathscr{H}$ Hilbert, but the theory described above carries over immediately to this situation. In particular, the inner product $\langle \cdot ; \cdot \rangle_{\mathscr{H}^2}$ and the norm $\| \cdot \|_{\mathscr{H}^2}$ in $\mathscr{H}^2$ are given by

$$\langle (x,y);(\tilde{x},\tilde{y})\rangle_{\mathscr{H}^2} = \langle x;\tilde{x}\rangle_{\mathscr{H}} + \langle y;\tilde{y}\rangle_{\mathscr{H}},$$

$$\|(x,y)\|_{\mathscr{H}^2}^2 = \|x\|_{\mathscr{H}}^2 + \|y\|_{\mathscr{H}}^2; \quad (x,y),(\tilde{x},\tilde{y}) \in \mathscr{H}^2,$$

where $\langle \cdot ; \cdot \rangle_{\mathscr{H}}$ and $\| \cdot \|_{\mathscr{H}}$ denote the inner product and norm in $\mathscr{H}$. Hence if (φ_m) is a CONS in $\mathscr{H}$, then $((\varphi_m, 0), (0, \varphi_m))$ is a CONS in $\mathscr{H}^2$.

By a slight abuse of notation, we will denote the inner product and norm in $\mathscr{H}_k^2$ by $\langle \cdot ; \cdot \rangle_k$ and $\| \cdot \|_k$, respectively. When $k = 0$, the subindex will be omitted.

3. Description of the models

3.1. Deterministic model

Consider a closed population consisting of susceptible, infective and immune individuals. The individuals live in a Euclidean space $\mathscr{V}$ and the dynamics of the model depends on their positions in the space. Thus, we introduce the functions $\bar{x}(r,t), \bar{y}(r,t)$ and $\bar{z}(r,t)$, denoting the density of susceptibles, infectives and immunes, respectively, at the position $r \in \mathscr{V}$ at time $t \ge 0$. Since the population is closed, we do not have to keep track of the immunes separately.

A given susceptible at position r may get infected by a given infective at another position r' with rate $\lambda(r, r')$. Thus the individuals do *not* move around, but infection occurs "at a distance". We may also assume that a given infective gets immune with a constant rate that is the same (equal to one, say) all over the space.

For simplicity, we assume that $\mathscr{V} = [0, 1]$. The above considerations lead us to the following system of PDEs:

$$\frac{\partial \bar{x}}{\partial t} = -\bar{x}\Lambda\bar{y}, \tag{3.1a}$$

$$\frac{\partial \bar{y}}{\partial t} = \bar{x}\Lambda\bar{y} - \bar{y},$$

with initial conditions

$$\bar{x}(r,0) = \bar{x}_0(r), \tag{3.1b}$$

$$\bar{y}(r,0) = \bar{y}_0(r),$$

where $\bar{x}_0, \bar{y}_0 \in C(\mathscr{V})$ and

$$\Lambda\bar{y}(r,t) = \int_0^1 \lambda(r,r')\bar{y}(r',t)\mathrm{d}r'. \tag{3.2}$$

For technical reasons we assume that the (measurable) rate function λ fulfills the following condition:

$$\rho^2 = \sup_{r\in[0,1]} \int_0^1 \lambda^2(r,r')\mathrm{d}r' < \infty, \tag{A1}$$

so that $\|\Lambda\bar{y}\|_\infty \le \rho\|\bar{y}\|$ ($\|\cdot\|_\infty$ denotes the supremum norm). For fixed $r \in \mathscr{V}$, the integral above may be regarded as a measure of the total infectiousness that a susceptible at r experiences, if the density of infectives is constant in space. Now, in order to see that system (3.1) has a *unique nonnegative continuous* solution $\bar{\xi} = (\bar{x}, \bar{y})$ on $\mathscr{V} \times [0, \infty)$, we consider instead the density $\bar{z}(r,t)$ of immune individuals. By the same reasoning as in Aronson (1977), Eq. (3.1) is equivalent to the following nonlinear equation:

$$\frac{\partial \bar{z}}{\partial t} = F(\bar{z}),$$

$$\bar{z}(r,0) = 0,$$

where

$$F(\bar{z}) = \bar{x}_0(1 - \mathrm{e}^{-\Lambda\bar{z}}) + \bar{y}_0 - \bar{z}.$$

The restriction of F to the set of nonnegative continuous functions is Lipschitz-continuous in the supremum norm, since

$$\|\Lambda z\|_\infty = \sup_{r\in[0,1]} \int_0^1 \lambda(r,r')z(r')\mathrm{d}r'$$

$$\le \sup_{r\in[0,1]} \int_0^1 \lambda(r,r')\mathrm{d}r'\,\|z\|_\infty \le \rho\|z\|_\infty$$

by (A1). We may now apply the method of successive approximations to get the result (cf. Theorem 2.2.1 in Lakshmikantham and Leela, 1981). For the nonnegativity of $\bar{x}$ and $\bar{y}$, see Aronson (1977).

3.2. Stochastic model

Now, we want to "justify" the deterministic model above, by recovering it as a limit of a sequence of stochastic multitype epidemics as the number of types, and the number of individuals of each type, both grow to infinity. Therefore, we may divide the space $\mathscr{V} = [0,1]$ into N groups, each group containing approximately ℓ individuals. Then, setting $n = (N,\ell)$, we consider, for $t \geq 0$,

$$\tilde{\boldsymbol{X}}^n(t) = (\tilde{X}_1^n(t), \dots, \tilde{X}_N^n(t)) \quad \text{and} \quad \tilde{\boldsymbol{Y}}^n(t) = (\tilde{Y}_1^n(t), \dots, \tilde{Y}_N^n(t)),$$

where

$$\tilde{X}_i^n(t) = \#\{\text{susceptibles in the } i\text{th group at time } t\}$$

and

$$\tilde{Y}_i^n(t) = \#\{\text{infectives in the } i\text{th group at time } t\}.$$

To obtain the classical multitype general epidemic, we model the population density $(\boldsymbol{X}^n, \boldsymbol{Y}^n) = (\tilde{\boldsymbol{X}}^n/\ell, \tilde{\boldsymbol{Y}}^n/\ell)$ as a Markov jump process with state space $\{(\boldsymbol{k}/\ell, \boldsymbol{m}/\ell);\ \boldsymbol{k},\boldsymbol{m} \in \mathbf{N}^N\}$ and with jump intensities of the following form:

$$q\left(\left(\frac{\boldsymbol{k}}{\ell},\frac{\boldsymbol{m}}{\ell}\right);\left(\frac{\boldsymbol{k}}{\ell}-\frac{\boldsymbol{e}^i}{\ell},\frac{\boldsymbol{m}}{\ell}+\frac{\boldsymbol{e}^i}{\ell}\right)\right) = \frac{\ell}{N}\left(\frac{k_i}{\ell}\right)\sum_{j=1}^{N}\lambda_{ij}^N\left(\frac{m_j}{\ell}\right),$$

$$q\left(\left(\frac{\boldsymbol{k}}{\ell},\frac{\boldsymbol{m}}{\ell}\right);\left(\frac{\boldsymbol{k}}{\ell},\frac{\boldsymbol{m}}{\ell}-\frac{\boldsymbol{e}^i}{\ell}\right)\right) = \ell\left(\frac{m_i}{\ell}\right), \quad 1 \leq i \leq N, \tag{3.3}$$

$$q = 0 \quad \text{otherwise},$$

where the numbers λ_{ij}^N are nonnegative, and where $\boldsymbol{e}^i$ denotes the ith unit vector in $\mathbf{R}^N$. (All the processes $(\boldsymbol{X}^n, \boldsymbol{Y}^n)$ are defined on a common probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$.)

The scaling by $1/N$ in front of the sum above reflects the fact that the (long range) interaction between the individuals is *weak*, and the ℓ-scaling comes from the classical density dependence assumption (cf. Kurtz, 1981). From now on, we identify every N-vector $\boldsymbol{x} = (x_1, \dots, x_N)$ with the step function $x = \sum_{i=1}^N x_i h^i \in \mathscr{S}_N$. In this way, our stochastic process becomes a space–time jump Markov process

$$\Xi^n(r,t) = (X^n(r,t), Y^n(r,t)), \quad (r,t) \in \mathscr{V} \times [0, \infty),$$

with state space $\mathscr{S}_N^2$ and infinitesimal generator A_n, given by

$$A_n f(x,y) = \ell N\left\langle x(\cdot)\Lambda_N y(\cdot); f\left(x - \frac{h^{\cdot}}{\ell}, y + \frac{h^{\cdot}}{\ell}\right) - f(x,y)\right\rangle$$

$$+ \ell N\left\langle y(\cdot); f\left(x, y - \frac{h^{\cdot}}{\ell}\right) - f(x,y)\right\rangle, \tag{3.4}$$

where

$$\Lambda_N y = \sum_{i=1}^{N} \left\{ \frac{1}{N} \sum_{j=1}^{N} \lambda_{ij}^N y_j \right\} h^i. \tag{3.5}$$

We make the following hypothesis:

$$\|\Lambda_N y - \Lambda y\|_\infty \le \sigma_N \|y\| \quad \text{for all } y \in \mathscr{S}_N, \tag{A2}$$

where $\sigma_N \to 0$ as $N \to \infty$.

By the Hahn–Banach theorem, the operator $\Lambda_N - \Lambda$ may be extended to $\mathscr{H}_0$ in such a way that Condition (A2) still holds. A typical example is the case where $\Lambda_N y = P_N(\Lambda y)$. If, for example,

$$\Gamma^2 = \sup_{r \in [0,1]} \int_0^1 |D_1^1 \lambda(r, r')|^2 \, \mathrm{d}r' < \infty,$$

then, by (2.5),

$$\|\Lambda_N y - \Lambda y\|_\infty \le \sigma_N \|y\| \quad \text{for all } y \in \mathscr{H}_0,$$

where $\sigma_N = \Gamma/N$.

Define the discrepancy functions

$$\tilde{\lambda}_i^N(r) = \sum_{j=1}^{N} \tilde{\lambda}_{ij}^N h^j(r), \quad 1 \le i \le N,$$

where

$$\tilde{\lambda}_{ij}^N = \lambda_{ij}^N - N^2 \int_{I_i \times I_j} \lambda(r, r') \, \mathrm{d}r \, \mathrm{d}r',$$

and put

$$\psi^N(r) = \sum_{i=1}^{N} \|\tilde{\lambda}_i^N\| \, h^i(r).$$

Then, for $y \in \mathscr{S}_N$,

$$(\Lambda_N - P_N \Lambda) y(r) = \sum_{i=1}^{N} \langle \tilde{\lambda}_i^N ; y \rangle h^i(r)$$

satisfies

$$|(\Lambda_N - P_N \Lambda) y(r)| \le |\psi^N(r)| \|y\|. \tag{3.6}$$

Therefore, in view of Eq. (2.5), a triangle inequality yields

$$\sigma_N \le \|\psi^N\|_\infty + \Gamma/N \quad \text{or} \quad \|\psi^N\|_\infty \le \sigma_N + \Gamma/N.$$

This shows that (A2) is equivalent to the assumption that $\|\psi^N\|_\infty \to 0$, as $N \to \infty$.

Here is another particular case.

Example. Let us consider the case where the infection rate $\lambda(r, r')$ only depends on the relative distance $|r - r'|$. In order not to get problems at the boundary of $\mathscr{V} = [0, 1]$,

we impose periodic boundary conditions on the initial value $\bar{\xi}(0)$. Take an even function defined on $[-\frac{1}{2},\frac{1}{2}]$ and extend it periodically to the whole line to get our rate function λ.
Then

$$\Lambda\bar{y}(r,t) = \lambda * \bar{y}(r,t) = \int_0^1 \lambda(r-r')\bar{y}(r',t)\,\mathrm{d}r'.$$

By translation invariance, Condition (A1) is equivalent to the condition $\lambda \in \mathscr{H}_0$. In practice, one first estimates the approximating numbers λ_{ij}^N and then takes the limit to obtain the function λ. Then (A2) may or may not hold. On the other hand, if we assume that λ is given, we can derive its conditional expectations by putting

$$\lambda_{ij}^N = N\int_{I_{i-j}} \lambda(r)\,\mathrm{d}r, \quad 1 \le i,j \le N,$$

where, for given $N \ge 1, I_k = ((k-1)/N, k/N]$ for all integers k. Now, if λ is continuous, then a simple calculation yields that, for every $\varepsilon > 0$, there exists N_0 such that, for every $N \ge N_0$,

$$\|\Lambda_N y - \Lambda y\|_\infty \le \varepsilon\|y\|.$$

Hence, Condition (A2) certainly holds.

We will also require that

$$\sup_n E\|\Xi^n(0)\|^2 < \infty. \tag{A3}$$

To see that the limiting process of $\Xi^n(t)$ (if it exists!) should be continuous, note that, by (3.4), the jumps $\Delta\Xi^n(r,t) = \Xi^n(r,t) - \Xi^n(r,t^-)$ at time t satisfy

$$\sup_t \|\Delta\Xi^n(t)\| \le \frac{2}{\ell\sqrt{N}}, \tag{3.7}$$

which vanishes in the limit $N,\ell \to \infty$. More precisely, we have the following lemma.

Lemma 3.1. *Let* $\varepsilon_n = 1/(N\ell)$. *Then*

$$\sup_n \sup_t \left\|\frac{1}{\sqrt{\varepsilon_n}}\Delta\Xi^n(t)\right\| < \infty$$

and

$$\sup_t \left\|\frac{1}{\sqrt{\varepsilon_n}}\Delta\Xi^n(t)\right\| \to 0 \quad \text{a.s. } [\boldsymbol{P}],$$

as $N,\ell \to \infty$.

4. Law of large numbers

Let $\mathscr{F}^n = (\mathscr{F}^n_t)_{t\geq 0}$ denote the natural filtration of Ξ^n, and define the following operator:

$$b^n(\xi) = (-xA_N y, xA_N y - y); \quad \xi = (x, y) \in \mathscr{S}^2_N.$$

Applying Dynkin's formula to the coordinate functions $f^i(x,y) = x_i$ and $g^i(x,y) = y_i$ for $1 \leq i \leq N$, one sees that our process $\Xi^n = (\Xi^n(t), t \geq 0)$ (we suppress the space variable r) admits the following semimartingale representation:

$$\Xi^n(t) = \Xi^n(0) + \int_0^t b^n(\Xi^n(s))\mathrm{d}s + M^n(t), \quad t \geq 0, \tag{4.1}$$

where M^n is a local $\mathscr{F}^n$-martingale with values in $\mathscr{S}^2_N \subseteq \mathscr{H}^2_0$. Hence we may use the theory of Hilbert-valued martingales, as developed e.g. in Métivier (1982).

We denote by $\ll M^n \gg$ the unique $\mathscr{F}^n$-predictable $\mathscr{L}_1(\mathscr{H}^2_0, \mathscr{H}^2_0)$-valued process with the following property: For every $\varphi, \psi \in \mathscr{H}^2_0$, the process

$$(\langle M^n(t), \varphi\rangle\langle M^n(t), \psi\rangle - \langle \varphi, \ll M^n \gg_t \psi\rangle, t \geq 0)$$

is a local martingale. Moreover, if $\prec M^n \succ_t$ is defined to be the trace of $\ll M^n \gg_t$, then the process $\|M^n\|^2 - \prec M^n \succ$ is also a local martingale. Our process $\ll M^n \gg$ is described in the following proposition.

Proposition 4.1. *Let $\varepsilon_n = 1/(\ell N)$. Then the restriction of the nuclear operator $\ll M^n \gg_t$ to the space $\mathscr{S}^2_N$ is given by $\varepsilon_n \int_0^t a^n(\Xi^n(s))\mathrm{d}s$, where $a^n(\xi)$, $\xi = (x,y) \in \mathscr{S}^2_N$, is a $(2N \times 2N)$-matrix consisting of four diagonal blocks,*

$$a^n(\xi) = \begin{pmatrix} \alpha^n(\xi) & \beta^n(\xi) \\ \beta^n(\xi) & \gamma^n(\xi) \end{pmatrix},$$

where, for $1 \leq i \leq N$,

$$\alpha^n_{ii}(\xi) = x_i(A_N y)_i,$$

$$\beta^n_{ii}(\xi) = -x_i(A_N y)_i,$$

$$\gamma^n_{ii}(\xi) = x_i(A_N y)_i + y_i.$$

Also, $\langle \varphi, \ll M^n \gg_t \psi\rangle = \langle P_N\varphi, \ll M^n \gg_t P_N\psi\rangle$ if $\varphi, \psi \in \mathscr{H}^2_0$.

Proof. Let us write

$$M^n = \sum_{i=1}^N M^n_i(h^i, 0) + \sum_{i=1}^N M^n_{N+i}(0, h^i)$$

and, for arbitrary $\varphi \in \mathscr{S}^2_N$,

$$\varphi = \sum_{i=1}^N \varphi_i(h^i, 0) + \sum_{i=1}^N \varphi_{N+i}(0, h^i).$$

Then we may actually use finite-dimensional theory to show that the process

$$\left(M_i^n(t)M_j^n(t) - \frac{1}{\ell}\int_0^t a_{ij}^n(\Xi^n(s))\,\mathrm{d}s, t \geq 0\right)$$

is a local martingale for each pair (i,j), $1 \leq i,j \leq 2N$. The first statement of the proposition now follows readily, since

$$\langle M^n(t), \varphi\rangle\langle M^n(t), \psi\rangle - \left\langle \varphi, \varepsilon_n \int_0^t a^n(\Xi^n(s))\,\mathrm{d}s\,\psi \right\rangle$$
$$= \frac{1}{N^2}\sum_{i,j=1}^{2N}\left(M_i^n(t)M_j^n(t) - \frac{1}{\ell}\int_0^t a_{ij}^n(\Xi^n(s))\,\mathrm{d}s\right)\varphi_i\psi_j,$$

for $\varphi, \psi \in \mathscr{S}_N^2$. The last statement is immediate, since $\langle M^n(t), \varphi\rangle = \langle M^n(t), P_N\varphi\rangle$ for every $\varphi \in \mathscr{H}_0^2$. □

Remark. Note that there is no temporal increase worth mentioning in the norm $\|\Xi^n\|$, since no new individuals are ever created. In fact, it is easily checked that $\|X^n(t)\|^2 \leq \|\Xi^n(0)\|^2$, $\|Y^n(t)\|^2 \leq 2\|\Xi^n(0)\|^2$ and then $\|\Xi^n(t)\|^2 \leq 3\|\Xi^n(0)\|^2$ for all $t \geq 0$. These estimates will be used frequently in the forthcoming calculations, without referring to this remark.

Corollary 4.2. $\|M^n(t)\| \to 0$ *in probability as* $N, \ell \to \infty$. *The convergence is uniform on compact time intervals* $[0, T]$.

Proof. Let $\delta > 0$. Then, by Doob's inequality for Hilbert-valued local martingales (see Theorem 23.4 of Métivier, 1982) and Proposition 4.1,

$$\begin{aligned} P\left(\sup_{0\leq t\leq T}\|M^n(t)\| > \delta\right) &\leq \frac{4}{\delta^2}E\|M^n(T)\|^2 = \frac{4}{\delta^2}E \,\langle\!\langle M^n \rangle\!\rangle_T \\ &= \frac{4\varepsilon_n}{\delta^2}\int_0^T \sum_{i=1}^{2N} E\{a_{ii}^n(\Xi^n(s))\}\,\mathrm{d}s \\ &\leq \frac{4\sqrt{2}T}{\delta^2\ell}\{2(\rho + \sigma_N)E\|\Xi^n(0)\|^2 + E\|\Xi^n(0)\|\}. \end{aligned}$$

This last quantity tends to zero as $N, \ell \to \infty$, by assumptions (A2) and (A3). □

We are now in a position to state and prove a law of large numbers for the sequence of processes Ξ^n.

Theorem 4.3. (LLN) *Assume that*
(i) $\|\Xi^n(0) - \xi(0)\| \to 0$ *in probability,*
(ii) *the hypotheses* (A1)–(A3) *hold.*

Then, for all $T > 0$,

$$\sup_{0\leq t\leq T}\|\Xi^n(t) - \xi(t)\| \to 0 \quad \textit{in probability,}$$

as $\ell \to \infty$ *and* $N \to \infty$.

Proof. For fixed n and $\delta > 0$, define the stopping time

$$\tau = \tau(n, \delta) = \inf\{t \ge 0: \|\Xi^n(t) - \bar{\xi}(t)\| \ge \delta\}$$

and put $\hat{\Xi}^n(t) = \Xi^n(t \wedge \tau)$ and $\hat{\bar{\xi}}(t) = \bar{\xi}(t \wedge \tau)$ for $t \ge 0$. It is enough to consider these stopped processes. Indeed, let $\varepsilon > 0$ and choose $\delta > \varepsilon$. Then

$$\left\{\sup_{0 \le t \le T} \|\Xi^n(t) - \bar{\xi}(t)\| \ge \varepsilon\right\} = \left\{\sup_{0 \le t \le T \wedge \tau} \|\Xi^n(t) - \bar{\xi}(t)\| \ge \varepsilon\right\}.$$

Taking into account Eq. (3.7) and using the fact that $\bar{\xi}$ is bounded, we may assume

$$\sup_{0 \le t \le T} \|\mathbf{1}_{\{\tau > 0\}} \hat{\Xi}^n(t)\| \le \delta_1 \quad \text{for some } \delta_1 \ge \delta.$$

Since $\hat{\Xi}^n$ is finite-dimensional,

$$\sup_{0 \le t \le T} \|\mathbf{1}_{\{\tau > 0\}} \hat{\Xi}^n(t)\|_\infty \le \delta_2^n,$$

which implies that $\hat{\Xi}^n(t)$ has a bounded total jump rate for every $t \in [0, T]$.

Now, define the following operator b:

$$b(\xi) = (-x\Lambda y, x\Lambda y - y); \quad \xi = (x, y) \in \mathscr{H}_0^2.$$

Then the integrated form of (3.1) may be written as

$$\bar{\xi}(t) = \bar{\xi}(0) + \int_0^t b(\bar{\xi}(s))\,\mathrm{d}s, \quad t \ge 0. \tag{4.2}$$

Let us compare (4.1) with (4.2). We have

$$\|\hat{\Xi}^n(t) - \hat{\bar{\xi}}(t)\| \le \|\Xi^n(0) - \bar{\xi}(0)\| + \int_0^t \|b^n(\hat{\Xi}^n(s)) - b(\hat{\bar{\xi}}(s))\|\,\mathrm{d}s + \|\hat{M}^n(t)\|. \tag{4.3}$$

First consider the integrand in (4.3). By repeated use of Minkowski's inequality,

$$\begin{aligned}\|b^n(\hat{\Xi}^n(s)) - b(\hat{\bar{\xi}}(s))\| &\le 2\|\hat{X}^n(s)\Lambda_N \hat{Y}^n(s) - \hat{x}(s)\Lambda\hat{y}(s)\| + \|\hat{Y}^n(s) - \hat{y}(s)\| \\ &\le R^n \|\hat{\Xi}^n(s) - \hat{\bar{\xi}}(s)\| + R_1^n,\end{aligned}$$

where

$$R^n = 2\rho\|\bar{\xi}(0)\|_\infty + 4(\rho + \sigma_N)\|\Xi^n(0)\| + 1,$$

$$R_1^n = 4\sigma_N\|\bar{\xi}(0)\|_\infty\|\Xi^n(0)\|.$$

Then, writing $R_2^n = \|\Xi^n(0) - \bar{\xi}(0)\|$ and $R_3^n(t) = \|\hat{M}^n(t)\|$, it follows that

$$\|\hat{\Xi}^n(t) - \hat{\bar{\xi}}(t)\| \le R_1^n t + R_2^n + R_3^n(t) + R^n \int_0^t \|\hat{\Xi}^n(s) - \hat{\bar{\xi}}(s)\|\,\mathrm{d}s.$$

Now apply Gronwall's lemma to get

$$\sup_{0 \le t \le T} \|\hat{\Xi}^n(t) - \hat{\bar{\xi}}(t)\| \le \left(R_1^n T + R_2^n + \sup_{0 \le t \le T} R_3^n(t)\right) \mathrm{e}^{R^n T}. \tag{4.4}$$

R_1^n and R_2^n both tend to zero in probability by Conditions (A2), (A3) and (i), and $\sup_{0 \le t \le T} R_3^n(t) \to 0$ in probability by Corollary 4.2. Moreover, R^n is bounded in probability (uniformly in n). Hence the right-hand side of (4.4) converges to zero. □

5. Central limit theorem

In this section we study the limiting distribution of the sequence of fluctuation processes

$$V^n = \frac{1}{\sqrt{\varepsilon_n}}(\Xi^n - \bar{\xi}),$$

defined on the stochastic bases $(\Omega, (\mathscr{F}_t^n), P^n)$, where P^n denotes the law of (the coordinate process) V^n in the Skorohod space $\Omega = D([0,T], \mathscr{H}_{-2}^2)$. We may perform a first-order Taylor expansion of the operator b around $\bar{\xi}$, using Eqs. (4.1)–(4.2), to get

$$V^n(t) = V^n(0) + \int_0^t B(s)V^n(s)\mathrm{d}s + \frac{1}{\sqrt{\varepsilon_n}}M^n(t) + \delta^n(t),$$

where B is the Fréchet derivative of b, and where we expect that

$$\sup_{0 \le t \le T} \|\delta^n(t)\|_{-2} \to 0$$

in probability. A Girsanov transformation will then reduce the study of the limiting behavior of V^n to considering the martingale term $\varepsilon_n^{-1/2}M^n$. This is our next objective.

5.1. Convergence of the local martingale

In view of Proposition 4.2 and the law of large numbers established in Corollary 4.2 for the sequence M^n of local martingales, we expect a central limit theorem for the sequence of normalized processes $\varepsilon_n^{-1/2}M^n$ to be valid. The limiting process turns out to be living in the Hilbert distribution space $\mathscr{H}_{-2}^2 \subseteq (\mathscr{E}')^2$; in fact, we will prove weak convergence in $D([0,T], \mathscr{H}_{-2}^2)$.

For fixed $t \ge 0$, define the following bilinear form on $\mathscr{H}_2^2$:

$$\langle \varphi, C(t)\psi \rangle = \int_0^t \langle \varphi, a(\bar{\xi}(s))\psi \rangle \mathrm{d}s, \varphi, \psi \in \mathscr{H}_2^2, \tag{5.1a}$$

where, for $\varphi = (\varphi_1, \varphi_2)$, $\psi = (\psi_1, \psi_2)$ and nonnegative $\xi = (x, y) \in \mathscr{H}_0^2$, $a(\xi)$ is defined by

$$\begin{aligned}\langle \varphi, a(\xi)\psi \rangle = {} & \langle \varphi_1; x\Lambda y\psi_1 \rangle - \langle \varphi_1; x\Lambda y\psi_2 \rangle \\ & - \langle \varphi_2; x\Lambda y\psi_1 \rangle + \langle \varphi_2; (x\Lambda y + y)\psi_2 \rangle. \end{aligned} \tag{5.1b}$$

Let us show that $a(\xi)$ is a nuclear operator from $\mathscr{H}_2^2$ to its dual $\mathscr{H}_{-2}^2$. For this, it is enough to show that the trace of $a(\xi)$ is finite. Define $(\varphi_m)_{m=1}^{\infty}$ to be a CONS in $\mathscr{H}_2^2$. By Cauchy–Schwarz inequality, Assumption (A1) and the inequality

$$\|\varphi\psi\| \leq \|\varphi\|_1\|\psi\|_1 \quad \text{for} \quad \varphi, \psi \in \mathscr{H}_2^2 \subseteq \mathscr{H}_1^2, \tag{5.2}$$

we get

$$\begin{aligned}\langle \varphi_m, a(\xi)\varphi_m\rangle &= 2\rho\|x\|\,\|y\|\,\|\varphi_m\|_1^2 + \|y\|\,\|\varphi_m\|_1^2\\ &\leq ((\rho+1)\|\xi\|^2+1)\,\|\varphi_m\|_1^2.\end{aligned}$$

It follows that

$$\sum_{m=1}^{\infty}\langle \varphi_m, a(\xi)\varphi_m\rangle \leq ((\rho+1)\|\xi\|^2+1)\sum_{m=1}^{\infty}\|\varphi_m\|_1^2,$$

which is finite, since the imbedding $\mathscr{H}_2^2 \to \mathscr{H}_1^2$ is Hilbert–Schmidt.

Similarly, if we extend $a^n(\cdot)$ from $\mathscr{S}_N^2$ to $\mathscr{H}_0^2$ by putting $a^n(\xi) = a^n(P_N\xi)$ for $\xi \in \mathscr{H}_0^2$, by the same reasoning as above, we may also deduce that $a^n(\xi) \in \mathscr{L}_1(\mathscr{H}_2^2, \mathscr{H}_{-2}^2)$.

Lemma 5.1. (i) *For each $\xi \in \mathscr{H}_0^2$ and $\varphi \in \mathscr{H}_2^2$,*

$$\lim_{n\to\infty}\langle \varphi, (a^n(\xi) - a(\xi))\varphi\rangle = 0. \tag{5.3}$$

(ii) *If $\varphi, \psi \in \mathscr{H}_2^2$,*

$$\lim_{n\to\infty} E^n\left[\int_0^T \left|\left\langle \varphi, (a^n(\hat{\Xi}^n(s)) - a^n(\hat{\bar{\xi}}(s)))\psi\right\rangle\right| \mathrm{d}s\right] = 0. \tag{5.4}$$

Proof. (i) Write $\xi = (x, y)$ and $\varphi = (\varphi_1, \varphi_2)$. The first term of $\langle \varphi, (a^n(\xi) - a(\xi))\varphi\rangle$ is

$$\langle \varphi_1, \{P_N x \Lambda_N P_N y - x\Lambda y\}\varphi_1\rangle.$$

By repeated use of the triangle inequality we conclude that this term tends to zero as $n \to \infty$, since $\|P_N z - z\| \to 0$ for all $z \in \mathscr{H}_0$ and since (A2) is in force. The other terms of the inner product are treated in the same way.

(ii) If we can show that

$$E^n\left[\int_0^T \|\hat{\Xi}^n(s) - \hat{\bar{\xi}}(s)\|^2\,\mathrm{d}s\right] \to 0 \quad \text{as } n \to \infty, \tag{5.5}$$

then (5.4) readily follows. But, since the integrand is bounded, the law of large numbers yields (5.5). □

Proposition 5.2. *Assume that $N, \ell \to \infty$ and let $T > 0$. Then the sequence $\varepsilon_n^{-1/2}M^n$ converges weakly to M in $D([0, T], \mathscr{H}_{-2}^2)$, where $M = (M(t),\ t \geq 0)$ is a continuous process with independent Gaussian increments and with covariance operator $C = (C(t),$ $t \geq 0)$ given by* (5.1).

If $\tilde{P}$ denotes the law of M in $C([0, T], \mathscr{H}_{-2}^2)$, then with

$$Q_t(\varphi) = \langle \varphi, a(\bar{\xi}(t))\varphi\rangle,$$

for every $f \in C_0^\infty(\mathbf{R})$ *and* $\varphi \in \mathscr{H}_2^2$

$$f(\langle M(t),\varphi\rangle) - \tfrac{1}{2}\int_0^t f''(\langle M(s),\varphi\rangle)Q_s(\varphi)\,\mathrm{d}s$$

is a $(\tilde{P},\mathscr{F}_t)$*-martingale.*

Proof. Applying Lemmas 3.1 and 5.1, the proof follows from Theorem 2.3 of Métivier (1984). □

5.2. Convergence of the whole process

In the rest of this section we will tacitly assume that the initial values of the solution of (3.1) are twice continuously differentiable; $\bar{x}_0, \bar{y}_0 \in C^2(\mathscr{V})$. By the same reasoning as at the end of Section 3.1, it will follow that $\bar{\xi} = (\bar{x},\bar{y})$ is twice continuously differentiable if the restriction of F to the set of nonnegative functions is Lipschitz in the C^2-norm, defined by

$$\|x\|_{C^2} = \max_{0\le l\le 2}\left\|\frac{\partial^l x}{\partial r^l}\right\|_\infty$$

(ordinary differentiation). This last fact follows readily by assuming

$$\rho_1^2 = \sup_{r\in[0,1]}\sum_{l=0}^{2}\int_0^1 |\partial_1^l \lambda(r,r')|^2\,\mathrm{d}r' < \infty. \tag{B1}$$

Writing $\tilde{M}^n = \varepsilon_n^{-1/2}M^n$, we have

$$V^n(t) = V^n(0) + \int_0^t B(s)V^n(s)\,\mathrm{d}s + \tilde{M}^n(t) + \delta^n(t), \tag{5.6}$$

where, for $V = (V_1, V_2)$,

$$B(t)V = (-(V_1\Lambda\bar{y}(t) + \bar{x}(t)\Lambda V_2), V_1\Lambda\bar{y}(t) + \bar{x}(t)\Lambda V_2 - V_2)$$

and $\delta^n = \delta_1^n + \delta_2^n$, with

$$\delta_1^n(t) = \frac{1}{\sqrt{\varepsilon_n}}\int_0^t [b(\Xi^n(s)) - b(\bar{\xi}(s)) - \sqrt{\varepsilon_n}B(s)V^n(s)]\,\mathrm{d}s,$$

$$\delta_2^n(t) = \frac{1}{\sqrt{\varepsilon_n}}\int_0^t [b^n(\Xi^n(s)) - b(\Xi^n(s))]\,\mathrm{d}s.$$

Note that the drift term B may be written in the more accessible form

$$\langle B(t)V,\varphi\rangle = \langle V, L(t)\varphi\rangle,$$

where

$$\begin{aligned} L(t)\varphi &= (\Phi_1(t),\Phi_2(t)) \\ &= ((\varphi_2-\varphi_1)\Lambda\bar{y}(t), \Lambda^*(\bar{x}(t)(\varphi_2-\varphi_1)) - \varphi_2). \end{aligned}$$

(The adjoint operator Λ^* of Λ is given by

$$\Lambda^*\varphi(r') = \int_0^1 \lambda(r,r')\varphi(r)\mathrm{d}r, \quad \varphi \in \mathscr{H}_2.)$$

We need a further restriction on the function λ, in order for the quantity $\langle V, L(t)\varphi\rangle$ to be well-defined. We make the following assumption:

$$\rho_2^2 = \sum_{l=0}^{2} \int_0^1 \int_0^1 |D_2^l \lambda(r,r')|^2 \mathrm{d}r\,\mathrm{d}r' < \infty. \tag{B2}$$

Conditions (B1) and (B2) ensure that the operator $B(t)$ is continuous on $\mathscr{H}^2_{-2}$. Indeed, (B1) implies that

$$\|\Phi_1(t)\|_2 \le 16\rho_1 \|\bar{y}(t)\|_\infty \|\varphi\|_2. \tag{5.7}$$

Likewise, using (B2), we get

$$\|\Phi_2(t)\|_2 \le (12\rho_2^2 \|\bar{x}(t)\|_\infty^2 + 2)^{1/2} \|\varphi\|_2. \tag{5.8}$$

Hence, $L(t)\varphi = (\Phi_1(t), \Phi_2(t))$ satisfies

$$\|L(t)\varphi\|_2 \le C(\bar{\xi}(t))\|\varphi\|_2, \tag{5.9}$$

where $C^2(\bar{\xi}(t)) = 2 + (2^8\rho_1^2 + 12\rho_2^2)\|\bar{\xi}(t)\|_\infty^2$.

Therefore, for $V \in \mathscr{H}^2_{-2}$,

$$|\langle B(t)V, \varphi\rangle| \le C(\bar{\xi}(t))\|V\|_{-2}\|\varphi\|_2. \tag{5.10}$$

In exactly the same way, we get

$$|\langle (B(t) - B(s))V, \varphi\rangle| \le (2^8\rho_1^2 + 12\rho_2^2)^{1/2}\|\bar{\xi}(t) - \bar{\xi}(s)\|_\infty \|V\|_{-2}\|\varphi\|_2, \tag{5.11}$$

showing that the mapping $t \to B(t)$ is continuous.

Now, we go back to Eq. (5.6). Let us assume that $\ell = \ell(N)$ satisfies

$$\frac{\ell}{N} \to \infty \quad \text{and} \quad \frac{\ell}{N^3} \to 0. \tag{B3}$$

We also require the following from the initial distribution:

$$\sup_n E\|\Xi^n(0)\| < \infty \tag{B4}$$

(which is (A3)), and

$$V^n(0) = \frac{1}{\sqrt{\varepsilon_n}}(\Xi^n(0) - \bar{\xi}(0)) \to V(0) \tag{B5}$$

in distribution in $\mathscr{H}^2_{-2}$, where $V(0)$ and the limiting process M are independent, and

$$\frac{1}{\sqrt{\varepsilon_n}}\|\Xi^n(0) - \bar{\xi}(0)\|^2 \to 0 \quad \text{in probability}. \tag{B6}$$

Under some conditions on ψ^N (cf. Eq. (3.6)), the next lemma tells us that the remainder δ^n converges to zero as it should.

Lemma 5.3. *Given a sequence* $\ell = \ell(N)$ *for which*
(i) *Conditions* (B1)–(B6) *hold,*
(ii) $\varepsilon_n^{-1/2}\|\psi^N\|_\infty^2$ *is uniformly bounded in* N,
(iii) $\varepsilon_n^{-1/2}\|\psi^N\|_{\mathscr{L}^1} \to 0$ *as* $N \to \infty$,

then

$$\sup_{0\le t\le T}\|\delta^n(t)\|_{-2} \to 0 \quad \textit{in probability},$$

as $N \to \infty$.

Remark. $\|\cdot\|_{\mathscr{L}^1}$ denotes the usual $\mathscr{L}^1$-norm: $\|x\|_{\mathscr{L}^1} = \int_0^1 |x(r)|\,\mathrm{d}r$. Observe that, in the case where $\Lambda_N = P_N\Lambda$, $\psi^N \equiv 0$. Also, note that if $\lambda \in C^2(\mathscr{V}\times\mathscr{V})$, then (B1) and (B2) certainly hold.

Proof. We have

$$\begin{aligned}\delta_1^n(t) &= \frac{1}{\sqrt{\varepsilon_n}}\int_0^t [b(\Xi^n(s)) - b(\bar{\xi}(s)) - \sqrt{\varepsilon_n}B(s)V^n(s)]\,\mathrm{d}s \\ &= \frac{1}{\sqrt{\varepsilon_n}}\int_0^t \varepsilon_n(-V_1^n(s)\Lambda V_2^n(s), V_1^n(s)\Lambda V_2^n(s))\,\mathrm{d}s.\end{aligned}$$

Hence, by (B1),

$$\begin{aligned}\|\delta_1^n(t)\| &\le \sqrt{2\varepsilon_n}\int_0^T \|V_1^n(s)\Lambda V_2^n(s)\|\,\mathrm{d}s \\ &\le \rho_1\sqrt{2\varepsilon_n}\int_0^T \|V_1^n(s)\|\,\|V_2^n(s)\|\,\mathrm{d}s \\ &\le \frac{\rho_1}{\sqrt{2\varepsilon_n}}\int_0^T \|\Xi^n(s) - \bar{\xi}(s)\|^2\,\mathrm{d}s,\end{aligned}$$

which implies that

$$\sup_{0\le t\le T}\|\delta_1^n(t)\| \le \frac{\rho_1 T}{\sqrt{2\varepsilon_n}}\sup_{0\le t\le T}\|\Xi^n(t) - \bar{\xi}(t)\|^2. \tag{5.12}$$

To see that the right-hand side of (5.12) converges to zero in probability, the argument used in the beginning of the proof of Theorem 4.3 shows that it suffices to consider the stopped process. Now, using the fact that $\sigma_N \le \|\psi^N\|_\infty + \rho_1/N$ and assumptions (B3), (B6) and (ii) the assertion follows from Eq. (4.4).

The term δ_2^n is estimated as follows:

$$\sup_{0\le t\le T} \|\delta_2^n(t)\|_{-2} \le \frac{1}{\sqrt{\varepsilon_n}} \int_0^T \|b^n(\Xi^n(s)) - b(\Xi^n(s))\|_{-2}\,\mathrm{d}s$$

$$\le \frac{2T}{\sqrt{\varepsilon_n}} \sup_{0\le t\le T} \|X^n(t)(\Lambda_N - \Lambda)Y^n(t)\|_{-2}.$$

Consider

$$\frac{1}{\sqrt{\varepsilon_n}} \|X^n(t)(\Lambda_N - \Lambda)Y^n(t)\|_{-2} \le \sqrt{\ell N}\,\|(X^n(t) - \bar{x}(t))(\Lambda_N - \Lambda)Y^n(t)\|_{-2}$$

$$+ \sqrt{\ell N}\,\|\bar{x}(t)(\Lambda_N - \Lambda)Y^n(t)\|_{-2}$$

$$= \mathrm{I} + \mathrm{II}.$$

To estimate I, we observe that, by (B1) and Eq. (2.5),

$$\|(P_N - I)\Lambda Y^n(t)\|_\infty \le \frac{\rho_1}{N}\|Y^n(t)\|.$$

Hence, using Eq. (3.6),

$$I \le \sqrt{\ell N}\,\|(X^n(t) - \bar{x}(t))(\Lambda_N - P_N\Lambda)Y^n(t)\|_{-2}$$

$$+ \sqrt{\ell N}\,\|(X^n(t) - \bar{x}(t))(P_N\Lambda - \Lambda)Y^n(t)\|_{-2}$$

$$\le \sqrt{\ell N}\,\|X^n(t) - \bar{x}(t)\|\,\|\psi^N\|_\infty\|Y^n(t)\| + \sqrt{\ell N}\,\|X^n(t) - \bar{x}(t)\|\frac{\rho_1}{N}\|Y^n(t)\|.$$

These expressions both converge to zero in probability (uniformly in t) by (ii), (B3) and the argument given after (5.12).

Now consider the term II above. Note that $\bar{x}(t)$ defines a bounded linear operator through multiplication on $\mathscr{H}_{-2}$. Indeed, by duality, we only need to show that $\|\bar{x}(t)\varphi\|_2 \le K\|\varphi\|_2$ if $\varphi \in \mathscr{H}_2$. But this is immediate, since $\bar{x}(t) \in C^2(\mathscr{V})$.

Noting that $\|f\|_{-2} \le C_1\|f\|_{\mathscr{L}^1}$ and applying Lemma 2.1, Eq. (3.6) and (B1), we get

$$\mathrm{II} \le K\sqrt{\ell N}(\|(\Lambda_N - P_N\Lambda)Y^n(t)\|_{-2} + \|(P_N\Lambda - \Lambda)Y^n(t)\|_{-2})$$

$$\le K\sqrt{\ell N}(C_1\|(\Lambda_N - P_N\Lambda)Y^n(t)\|_{\mathscr{L}^1} + \|(P_N\Lambda - \Lambda)Y^n(t)\|_{-2})$$

$$\le K\sqrt{\ell N}\left(C_1\|\psi^N\|_{\mathscr{L}^1} + \frac{\rho_1 C_2}{N^2}\right)\|Y^n(t)\|.$$

It follows from (B3), (B4) and (iii) that II converges to zero in probability, uniformly in t. □

In view of Proposition 5.2, Condition (B5) and Lemma 5.3, the semimartingale representation (5.6) of V^n suggests that the limiting process V (if it exists!) should be in $C([0,T],\mathscr{H}^2_{-2})$ and would be the mild solution of the following equation:

$$V(t) = V(0) + \int_0^t B(s)V(s)\,\mathrm{d}s + M(t). \tag{5.13}$$

The next theorem shows indeed that V^n converges weakly in $D([0,T],\mathscr{H}^2_{-2})$ to the solution of Eq. (5.13). (The present proof, which replaces an earlier approach of our own, is due to D. Blount.)

Theorem 5.4 (CLT). *Assume that the conditions of Lemma 5.3 hold. Then V^n converges weakly in $D([0,T],\mathscr{H}^2_{-2})$ to the limit $V \in C([0,T],\mathscr{H}^2_{-2})$ which is the mild solution of Eq.* (5.13).

Proof. Let $U(t,s)$ be the operator-valued solution of

$$\frac{\partial}{\partial t}U(t,s) = B(t)U(t,s), \quad 0 \le s \le t \le T,$$

$$U(t,t) = I. \tag{5.14}$$

In view of Eqs. (5.10) and (5.11), Theorem 5.2 in Pazy (1983) ensures the existence of such a solution, for which $(t,s) \to U(t,s)$, $0 \le s \le t \le T$, is continuous in the uniform operator topology.

Now, the continuity of U and B yields that

$$f(\alpha,\beta) = U(\cdot,0)\alpha + \beta(\cdot) + \int_0^{\cdot} U(\cdot,s)B(s)\beta(s)\,\mathrm{d}s$$

defines a continuous function from $\mathscr{H}^2_{-2} \times D([0,T],\mathscr{H}^2_{-2})$ into $D([0,T],\mathscr{H}^2_{-2})$. V^n admits the following representation:

$$V^n(t) = U(t,0)V^n(0) + (\tilde{M}^n(t) + \delta^n(t))$$
$$+ \int_0^t U(t,s)B(s)(\tilde{M}^n(s) + \delta^n(s))\,\mathrm{d}s. \tag{5.15}$$

To see this, note that by Gronwall's inequality, Eq. (5.6) has a unique solution. The claim follows them by simply computing $\mathrm{d}V^n$ from (5.15) using (5.14). Therefore V^n can be written as

$$V^n = f(V^n(0), \tilde{M}^n + \delta^n).$$

Using Proposition 5.2, Condition (B5) and Lemma 5.3, it follows that

$$(V^n(0), \tilde{M}^n + \delta^n) \to (V(0), M)$$

in distribution in $\mathscr{H}^2_{-2} \times D([0,T],\mathscr{H}^2_{-2})$. The theorem now follows from the continuous mapping theorem, giving the limiting process $V = f(V(0),M)$ that satisfies (5.13). □

Acknowledgement

We wish to thank the referee for carefully reading the paper, and D. Blount for many useful suggestions, in particular the present short form of the proof of the CLT.

References

R.A. Adams, Sobolev Spaces (Academic Press, New York, 1975).
L. Arnold, Mathematical models of chemical reactions, in: M. Hazewinkel and J. Willems, eds., Stochastic Systems (Reidel, Dordrecht, 1981).
D.G. Aronson, The asymptotic speed of propagation of a simple epidemic, in: W.E. Fitzgibbon (III) and H.F. Walker, eds., Nonlinear Diffusion (Pitman, London, 1977)..
D. Blount, Comparison of stochastic and deterministic models of a linear chemical reaction with diffusion, Ann. Probab. 19 (1991) 1440–1462.
D. Blount, Limit theorems for a sequence of nonlinear reaction-diffusion systems, Stochastic Process. Appl. 45 (1993) 193–207.
S.N. Ethier and T.G. Kurtz, Markov Processes, Characterization and Convergence (Wiley, New York, 1986).
I.M. Gelfand and N.Y. Vilenkin, Generalized Functions, Vol. IV (Academic Press, New York, 1964).
R.A. Holley and D.W. Stroock, Generalized Ornstein–Uhlenbeck processes and infinite particle branching Brownian motions, Publ. RIMS Kyoto Univ. 14 (1978) 741–788.
D.G. Kendall, Mathematical models of the spread of infections, Mathematics and Computer Science in Biology and Medicine, Medical Research Council (1965).
P. Kotelenez, Law of large numbers and central limit theorem for linear chemical reactions with diffusion, Ann. Probab. 14 (1986) 173–193.
P. Kotelenez, High density limit theorems for nonlinear chemical reactions with diffusion, Probab. Theory Related Fields 78 (1988) 11–37.
T.G. Kurtz, Limit theorems for sequences of jump Markov processes approximating ordinary differential processes, J. Appl. Probab. 9 (1971) 344–356.
T.G. Kurtz, Approximation of population processes, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 36 (SIAM, Philadelphia, 1981).
V. Lakshmikantham and S. Leela, Nonlinear Differential Equations in Abstract Spaces (Pergamon, Oxford, 1981).
A. Martin-Löf, Limit theorems for the motion of a Poisson system of independent Markovian particles with high density, Z. Wahrscheinlichskeitheor. Verw. Geb. 34 (1976) 205–223.
M. Métivier, Semimartingales (De Gruyter, Berlin, 1982).
M. Métivier, Convergence faible et principe d'invariance pour des martingales à valeurs dans des espaces de Sobolev, Ann. Inst. Henri Poincaré 20 (1984) 329–348.
I. Mitoma, Tightness of probabilities on $C([0,1];\mathscr{S}')$ and $D([0,1];\mathscr{S}')$, Ann. Probab. 11 (1983) 989–999.
K. Oelschläger, On the derivation of reaction–diffusion equations as limit dynamics of systems of moderately interacting stochastic processes, Probab. Theory Related Fields 82 (1989) 565–586.
A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations (Springer, New York, 1983).